



HAL
open science

Génération de la “ banane de la parole ” en vue d’une évaluation objective de l’intelligibilité

Timothy Pommée, Julie Mauclair, Virginie Woisard, Jérôme Farinas, Julien Pinquier

► To cite this version:

Timothy Pommée, Julie Mauclair, Virginie Woisard, Jérôme Farinas, Julien Pinquier. Génération de la “ banane de la parole ” en vue d’une évaluation objective de l’intelligibilité. 8e Journées de Phonétique Clinique (JPC 2019), May 2019, Mons, Belgique. pp.107-109. hal-03631165

HAL Id: hal-03631165

<https://hal.science/hal-03631165>

Submitted on 5 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive Toulouse Archive Ouverte

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in: <http://oatao.univ-toulouse.fr/24885>

To cite this version: Pommée, Timothy and Mauclair, Julie and Woisard, Virginie and Farinas, Jérôme and Pinquier, Julien *Génération de la « banane de la parole » en vue d'une évaluation objective de l'intelligibilité.* (2019) In: 8e Journées de Phonétique Clinique (JPC 2019), 14 May 2019 - 16 May 2019 (Mons, Belgium)

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

Génération de la « banane de la parole » en vue d'une évaluation objective de l'intelligibilité

Timothy POMMÉE¹

Julie MAUCLAIR¹

Virginie WOISARD^{2,3}

Jérôme FARINAS¹

Julien PINQUIER¹

¹ Institut de Recherche en Informatique de Toulouse,
CNRS, Université Toulouse III, France

² Oncopole, Toulouse, France

³ Laboratoire Octogone Lordat, Toulouse, France

L'aire du triangle vocalique, construit en mesurant les valeurs fréquentielles des deux premiers formants des voyelles prononcées, est l'une des mesures employées pour l'évaluation de l'intelligibilité de la parole [1,2]. Le placement des voyelles sur un graphique à deux axes, F1 et F2, permet d'interpréter les valeurs obtenues par rapport au triangle « standard ». Toutefois, les consonnes ont une importance cruciale dans l'intelligibilité de la parole. Stevens et Blumstein [3] affirmaient déjà en 1978 que le lieu d'articulation des consonnes occlusives serait identifiable par l'aspect global de leur représentation spectrale. En audiométrie, la « banane de la parole » est la représentation de la composante fréquentielle principale « standard » de chaque consonne, ainsi que de son intensité, permettant la délimitation d'une aire sur un audiogramme. Cette démarche a été réalisée pour les consonnes de l'anglais [4,5] et du Thaï [6]. Pour le français, Béchet et al. [7] ont calculé des aires consonantiques sur les occlusives sonores [b,d,g], représentant les trois lieux d'articulation les plus communs à l'échelle universelle [8], à l'aide des F2 et F3. Ainsi, nous avons tenté de construire la banane de la parole sur la base des 16 consonnes principales du français.

Deux sujets sains (1 homme, 1 femme) ont produit ces 16 consonnes entourées des voyelles « extrêmes » [a,i,u] et des voyelles plus neutres [œ,ø], à trois reprises. En nous inspirant des travaux de Klangpornkun [6], nous avons employé le codage prédictif linéaire (LPC) pour identifier les pics spectraux proéminents des consonnes [9,10] dans les différents contextes vocaliques, et selon le genre du locuteur. En positionnant ces consonnes sur un graphique à deux axes – fréquence et

intensité, nous avons ensuite généré la banane de la parole des consonnes du français.

Notons qu'au vu de la variabilité interindividuelle des données acoustiques [11,12,13,14], le nombre de sujets dans cette étude préliminaire est limité. De nouveaux enregistrements dans des conditions standardisées sont en cours, visant à obtenir des extraits de parole d'au moins 2 hommes et 2 femmes des tranches d'âge 20-29, 30-39, 40-49, 50-59, 60-69 et 70+.

Par l'automatisation de la construction d'une telle représentation des consonnes produites par le sujet, nous espérons pouvoir proposer un outil d'évaluation objective de l'intelligibilité de la parole.

Références bibliographiques

- [1] WEISMER, G., LAURES, J., JENG, J.-Y., KENT, R., KENT, J., Effect of speaking rate manipulations on acoustic and perceptual aspects of the dysarthria in amyotrophic lateral sclerosis, *Folia Phoniatrica et Logopaedica*, 2000, **Vol. 52/5**, 201-219. <https://doi.org/10.1159/000021536>
- [2] TURNER, G., TJADEN, K., WEISMER, G., The influence of speaking rate on vowel space and intelligibility for individuals with amyotrophic lateral sclerosis, *Journal of Speech & Hearing Research*, 1995, **Vol 38**, 1001-1013.
- [3] STEVENS, K.N., BLUMSTEIN, S.E., Invariant cues for place of articulation in stop consonants, *Journal of the Acoustical Society of America*, 1978, **Vol 64(5)**, 1358-1368.
- [4] NORTHERN, J.L., DOWNS, M.P., *Hearing in Children*, Lippincott Williams & Wilkins, 1984.
- [5] SKINNER, M., The hearing of speech during language acquisition. *Otolaryngol Clinics of North America*, 1978, **Vol 11**, 631-650.
- [6] KLANGPORNKUN, N., ONSUWAN, C., TANTIBUNDHIT, C., PITATHAWATCHAI, P., Predictions from « speech banana » and audiograms: Assessment of hearing deficits in Thai hearing loss patients. *Proceedings of Meetings on Acoustic*, 2013, **Vol 20**, 060004. <https://doi.org/10.1121/1.4879262>
- [7] BECHET, M., FERBACH-HECKER, V., HIRSCH, F., SOCK, R., F2/F3 of voiced plosives in VCV sequences in children with cleft palate: an acoustic study, *International Seminar on Speech Production*, Montréal, 2011, 65-73.
- [8] ABRY, C., [b]-[d]-[g] as a universal triangle as acoustically optimal as [i]-[a]-[u], *The 15th International Congress of Phonetic Sciences*, Barcelone, 2003, 727-730.
- [9] JACKSON, P.J.B., Acoustic cues of voiced and voiceless plosives for determining place of articulation, *CRAC workshop*, Aalborg, 2001, 19-22.

- [10] SERNICLAES, W., BOGLIOTTI, C., CARRE, R., Perception of consonant place of articulation: phonological categories meet natural boundaries, *15th ICPbS*, Barcelone, 2003, 391-394.
- [11] JOHNSON, K., LADEFOGED, P., LINDAU, M., Individual differences in vowel production, *Journal of the Acoustical Society of America*, 1993, **Vol 94(2)**, 701-714.
- [12] LAVOIE, L., Subphonemic and suballophonic consonant variation: The role of the phoneme inventory, *ZAS Papers in Linguistics*, 2002, **Vol 28**, 39-54.
- [13] FANT, G., *Speech sounds and features*, MIT Press, 1973.
- [14] PETERSON, G., BARNEY, H., Control methods used in a study of the vowels, *Journal of the Acoustical Society of America*, 1952, **Vol 24(2)**, 175-184.